



**The GfK Group Project Report
for the
Tobacco Products and Risk Perceptions Survey – 2015 (Year 2)**

Submitted to:
Deanne Weber
Porter Novelli

Date submitted: October 1, 2015

GfK PROJECT DIRECTOR:	Nicole Neuenschwander
GfK ACCOUNT EXECUTIVE:	Larry Osborn
GfK PROJECT NUMBER:	310.111.00545.1


GfK Deliverable Authorization			
Printed Name	Signature	Date	Title
Sergei Rodkin, Ph.D.		10/01/2015	VP, Government and Academic Research

Table of Contents

STUDY DESIGN & DOCUMENTATION	4
INTRODUCTION	4
SAMPLE DEFINITION	4
DATA COLLECTION FIELD PERIOD & SURVEY LENGTH	4
SURVEY COMPLETION AND SAMPLE SIZES	4
<i>Key Survey Response Statistics: In-Field Screening and Main Interview.....</i>	<i>4</i>
SURVEY COOPERATION ENHANCEMENTS.....	5
DATA FILE DELIVERABLES AND DESCRIPTIONS	5
<i>Supplemental Variables</i>	<i>7</i>
KEY PERSONNEL.....	11
GFK METHODOLOGY	12
INTRODUCTION	12
THE GFK GROUP.....	12
KNOWLEDGEPANEL METHODOLOGY INFORMATION	12
<i>ABS Recruitment.</i>	<i>13</i>
<i>Household Member Recruitment.....</i>	<i>14</i>
SURVEY SAMPLING FROM KNOWLEDGEPANEL	14
SURVEY ADMINISTRATION.....	14
SAMPLE WEIGHTING.....	15
<i>Study-Specific Post-Stratification Weights.....</i>	<i>15</i>
APPENDIX A: QUESTIONNAIRE	18
APPENDIX B: CODEBOOK.....	59
APPENDIX C: KNOWLEDGEPANEL® RESPONSE RATE REPORT	187

Study Design & Documentation

Introduction

The GfK Group (GfK, formerly Knowledge Networks) conducted the Longitudinal Smoker Survey – 2015 (Year 2) on behalf of Porter Novelli. Specifically, the study examines the awareness, attitudes, and beliefs of the general population and smokers on tobacco products. The survey was conducted using sample from KnowledgePanel®.

Sample Definition

The target population consists of the following: non-institutionalized adults age 18 and over residing in the United States and smokers (every day or some days). The sample excluded anyone that completed the Year 1 Smoker Survey.

To sample the population, GfK sampled households from its KnowledgePanel, a probability-based web panel designed to be representative of the United States.

To qualify for the general population group, a panel member must have been:

- 18+ years old;

To qualify for the smoker augment group, a panel member must have:

- Been 18+ years old;
- Reconfirmed smoker status (every day or some days), as determined by the survey questions

Data Collection Field Period & Survey Length

The data collection field periods were as follows:

Stage	Start Date	End Date
Pre-Test	08/06/2015	08/08/2015
Main	08/13/2015	09/21/2015

Participants completed the main survey in 25 minutes (median).

Survey Completion and Sample Sizes

The number of respondents sampled and participating in the survey, the survey completion rates for the screener and main interview, and the incidence/eligibility rate are presented below.

Key Survey Response Statistics: In-Field Screening and Main Interview

	N Sampled for Screener	N Complete Screener	Screener Survey Completion Rate	Qualified for Main Survey	Incidence Rate
General Population	7,194	5,498	76.4%	5,497	99.9%
Smoker Augment	941	683	72.6%	594	87.0%
Total	8,135	6,181	76.0%	6,091	98.5%

For the General Population group, screening refers to terminating respondents who refuse the first two survey questions about smoking. These two questions were crucial to weighting so respondents were disqualified for refusing.

For the Smoker Augment, screening refers to the reconfirmation of smoking status. Previously identified smokers who no longer smoked every day or some days (as determined in the survey) were disqualified.

Of the 6,091 cases completing the main survey, 6,051 cases were determined to be valid cases to be included in the final analyses. 38 cases were removed from the General Population group and 2 cases were removed from the Smoker Augment group. The 40 total cases were excluded due to refusing to answer more than one-half of the substantive survey questions. Any such exclusions were confirmed by the client.

Survey Cooperation Enhancements

As a standard, email reminders to non-responders were sent on day three of the field period.

Beyond the standard email reminder on day three of the field period, the following steps were also taken:

- Additional email reminders to non-responders were sent on day 7, 13, 21, and 29 of the field period;
- Participants received a cash-equivalent of \$5 for their participation. This increased to \$10 if they did not respond within the first 13 days of the field period.

Data File Deliverables and Descriptions

GfK prepared and delivered a fully formatted SPSS file containing the collected data, GfK demographic profile data, and the appropriate variable and value labels, as described below.

Data File Deliverables

Delivery Date	File Type	File Name	File Size
08/10/2015	SPSS	PorterNovelli_2015Smoker Survey_Prestest_Client.sav	454 KB
09/23/2015	SPSS	PorterNovelli_2015Smoker Survey_Main_Unweighted_Client.zip	1 Mb
09/28/2015	SPSS	PorterNovelli_2015Smoker Survey_Main_Weighted_Client.zip	1 Mb
09/29/2015	SPSS	PorterNovelli_SmokerSurvey_Y1 and Y2 pooled weights.zip	150 KB
10/01/2015	SPSS	PorterNovelli_2015Smoker Survey_Main_Weighted_Client.zip	2 Mb

In addition, GfK prepared and delivered other deliverables as follows:

- Post-stratification statistical weights for:
 - Total respondents, scaled to population and sample sizes

- General Population group, scaled to population and sample sizes,
- All Smokers, scaled to population and sample sizes, and
- Total respondents pooled from Year 1 and Year 2, scaled to population and sample sizes
- Open-ended Responses [NOT CODED]
- Demographic profile data for
 - All sampled panelists, including non-responders
- Additional non-demographic profile data for qualified respondents
 - Sexual Orientation (Q23), Do you consider yourself to be...
 - Self-reported health status (Q1), In general, would you say your physical health is...
 - Personal medical conditions (Q19), Have you been diagnosed with any of the following medical conditions?)
 - Sources of health information (Q51), Have you used any of the following sources for health information?
 - Political ideology
 - Early adopter
 - Internet sources of health/wellness information (Q52), Have you used any of the following Internet resources for health information?
 - ZIP
 - Device type
 - Number of smoking studies done in the past year (since August 2014)

Several supplemental variables are provided to assist the principal investigators in identifying cases that could potentially be of interest.

Please also note the following for the survey data file:

- When a respondent refused to answer a question, the code “-9” is used.

The table below shows the name and description of each of the supplemental, demographic, and other profile variables delivered to the client.

Supplemental Variables

Variable Name	Variable Description
CASEID	Case Identification Number
weight1	Post-Stratification weight: Gen Pop Respondents (sample based)
weight1pop	Post-Stratification weight: Gen Pop Respondents (population based)
weight2	Post-Stratification weight: Total Respondents (Gen Pop and Smoker Augment) weights (sample based)
weight2pop	Post-Stratification weight: Total Respondents (Gen Pop and Smoker Augment) weights (population based)
weight3	Post-Stratification weight: Total Smoker Respondents (Gen Pop and Pre-Id Smoker Augment) weights (sample based)
weight3pop	Post-Stratification weight: Total Smoker Respondents (Gen Pop and Pre-Id Smoker Augment) weights (population based)
weight4	Post-Stratification weight: Total Years 1 and 2 Respondents (Gen Pop + Pre-Id Smoker Augment) weights (sample based)
weight4pop	Post-Stratification weight: Total Years 1 and 2 Respondents (Gen Pop + Pre-Id Smoker Augment) weights (population based)
TM_START	Date and time interview started
TM_FINISH	Date and time interview ended
DURATION	Duration of interview in minutes
PPAGE	Age
PPAGECAT	Age - 7 Categories
PPAGECT4	Age - 4 Categories
PPEDUC	Education (Highest Degree Received)
PPEDUCAT	Education (Categorical)
PPETHM	Race / Ethnicity
PPGENDER	Gender
PPHHHEAD	Household Head
PPHHSIZE	Household Size
PPHOUSE	Housing Type
PPINCIMP	Household Income
PPMARIT	Marital Status
PPMSACAT	MSA Status
PPREG4	Region 4 - Based on State of Residence
PPREG9	Region 9 - Based on State of Residence
PPRENT	Ownership Status of Living Quarters
PPSTATEN	State
PPT01	Presence of Household Members - Children 0 - 2
PPT25	Presence of Household Members - Children 2 - 5
PPT612	Presence of Household Members - Children 6 - 12
PPT1317	Presence of Household Members - Children 13 - 17
PPT18OV	Presence of Household Members - Adults 18+
PPWORK	Current Employment Status
PPNET	Household Internet Access
ppp20063	Q23: Do you consider yourself to be . . .

pph10001	Q1: In general, would you say your physical health is. . .?
pph1acid	Q19: Have you been diagnosed with any of the following medical conditions? [Acid reflux disease]
pph1adhd	Q19: Have you been diagnosed with any of the following medical conditions? [ADHD or ADD]
pph1anxi	Q19: Have you been diagnosed with any of the following medical conditions? [Anxiety disorder]
pph1brea	Q19: Have you been diagnosed with any of the following medical conditions? [Asthma, chronic bronchitis or COPD]
pph1afib	Q19: Have you been diagnosed with any of the following medical conditions? [Atrial Fibrillation]
pph1bipo	Q19: Have you been diagnosed with any of the following medical conditions? [Bipolar Disorder]
pph1canc	Q19: Have you been diagnosed with any of the following medical conditions? [Cancer (all types except skin cancer)]
pph1pain	Q19: Have you been diagnosed with any of the following medical conditions? [Chronic pain (such as low back pain, neck pain, or fibromyalgia)]
pph1cysf	Q19: Have you been diagnosed with any of the following medical conditions? [Cystic Fibrosis]
pph1depr	Q19: Have you been diagnosed with any of the following medical conditions? [Depression]
pph1diab	Q19: Have you been diagnosed with any of the following medical conditions? [Diabetes]
pph1epil	Q19: Have you been diagnosed with any of the following medical conditions? [Epilepsy]
pph1eyed	Q19: Have you been diagnosed with any of the following medical conditions? [Eye Disease]
pph1gout	Q19: Have you been diagnosed with any of the following medical conditions? [Gout]
pph1hatk	Q19: Have you been diagnosed with any of the following medical conditions? [Heart attack]
pph1hprb	Q19: Have you been diagnosed with any of the following medical conditions? [Heart disease]
pph1hepc	Q19: Have you been diagnosed with any of the following medical conditions? [Hepatitis C]
pph1hype	Q19: Have you been diagnosed with any of the following medical conditions? [High blood pressure]
pph1chol	Q19: Have you been diagnosed with any of the following medical conditions? [High cholesterol]
pph1hiv	Q19: Have you been diagnosed with any of the following medical conditions? [HIV or AIDS]
pph1kidn	Q19: Have you been diagnosed with any of the following medical conditions? [Kidney disease]
pph1meno	Q19: Have you been diagnosed with any of the following medical conditions? [Menopause]

pph1mood	Q19: Have you been diagnosed with any of the following medical conditions? [Mood Disorder]
pph1ms	Q19: Have you been diagnosed with any of the following medical conditions? [Multiple sclerosis]
pph1osar	Q19: Have you been diagnosed with any of the following medical conditions? [Osteoarthritis, joint pain or inflammation]
pph1oste	Q19: Have you been diagnosed with any of the following medical conditions? [Osteoporosis or osteopenia]
pph1pmen	Q19: Have you been diagnosed with any of the following medical conditions? [Perimenopause]
pph1psor	Q19: Have you been diagnosed with any of the following medical conditions? [Psoriasis]
pph1arth	Q19: Have you been diagnosed with any of the following medical conditions? [Rheumatoid arthritis]
pph1alle	Q19: Have you been diagnosed with any of the following medical conditions? [Seasonal allergies]
pph1scha	Q19: Have you been diagnosed with any of the following medical conditions? [Schizoaffective Disorder]
pph1scph	Q19: Have you been diagnosed with any of the following medical conditions? [Schizophrenia]
pph1skin	Q19: Have you been diagnosed with any of the following medical conditions? [Skin cancer]
pph1inso	Q19: Have you been diagnosed with any of the following medical conditions? [Sleep disorders such as sleep apnea or insomnia]
pph1stro	Q19: Have you been diagnosed with any of the following medical conditions? [Stroke]
pph1ment	Q19: Have you been diagnosed with any of the following medical conditions? [Other Mental Health Condition]
pph1othr	Q19: Have you been diagnosed with any of the following medical conditions? [Something else]
pph1none	Q19: Have you been diagnosed with any of the following medical conditions? [None of these]
pph20032	Q51: Have you used any of the following sources for health information? [Doctor]
pph20033	Q51: Have you used any of the following sources for health information? [Pharmacist]
pph20034	Q51: Have you used any of the following sources for health information? [Nurse, nurse practitioner, or physician's assistant]
pph20035	Q51: Have you used any of the following sources for health information? [Relative, friend, or co-worker]
pph20036	Q51: Have you used any of the following sources for health information? [Someone you know who has a particular medical condition]
pph20037	Q51: Have you used any of the following sources for health information? [Disease association or patient support group]
pph20038	Q51: Have you used any of the following sources for health

	information? [Educational forum at a local clinic, hospital, community center]
pph20039	Q51: Have you used any of the following sources for health information? [Pharmaceutical company]
pph20040	Q51: Have you used any of the following sources for health information? [Health insurance company]
pph20041	Q51: Have you used any of the following sources for health information? [Newspapers or magazines]
pph20042	Q51: Have you used any of the following sources for health information? [Television]
pph20043	Q51: Have you used any of the following sources for health information? [The Internet]
pph20044	Q51: Have you used any of the following sources for health information? [Something else]
pph20045	Q51: Have you used any of the following sources for health information? [Have not looked for health information in the past 12 months]
pph20107	Q51: Have you used any of the following sources for health information? [Social Media]
pph20108	Q51: Have you used any of the following sources for health information? [Health Care App]
ppp10012	Q11: In general, do you think of yourself as...
ppadopt1	I usually try new products before other people do..Agree/Disagree
ppadopt2	I often try new brands because I like variety and get bored with the same old thing..Agree/Disagree
ppadopt3	When I shop I look for what is new..Agree/Disagree
ppadopt4	I like to be the first among my friends and family to try something new..Agree/Disagree
ppadopt5	I like to tell others about new brands or technology..Agree/Disagree
pph20046	Q52: Have you used any of the following Internet resources for health information? [General search engines]
pph20047	Q52: Have you used any of the following Internet resources for health information? [Health-related websites]
pph20048	Q52: Have you used any of the following Internet resources for health information? [Online communities or social networks]
pph20049	Q52: Have you used any of the following Internet resources for health information? [Video sharing sites]
pph20050	Q52: Have you used any of the following Internet resources for health information? [Pharmaceutical company websites]
pph20051	Q52: Have you used any of the following Internet resources for health information? [Websites for specific drugs]
pph20052	Q52: Have you used any of the following Internet resources for health information? [Medical tourism website]
pph20053	Q52: Have you used any of the following Internet resources for health information? [Hospital or clinic websites]
pph20054	Q52: Have you used any of the following Internet resources for

	health information? [Government websites]
pph20055	Q52: Have you used any of the following Internet resources for health information? [News sites]
pph20056	Q52: Have you used any of the following Internet resources for health information? [Somewhere else]
zip	Zip Code
DeviceType	DOV: Device Type
smkStudies	# of smoking studies completed since August 2014

Key Personnel

Key personnel on the study include:

Larry Osborn –Vice President, Consulting. L. Osborn is based in Chicago.

Phone number: (312) 281-2469

Email: larry.osborn@gfk.com

Nicole Neuenschwander – Senior Research Manager, Government & Academic Research. N.

Neuenschwander is based in Chicago.

Phone number: (312) 931-0103

Email: nicole.neuenschwander@gfk.com

Magda Lasota – Senior Research Associate, Government & Academic Research. M. Lasota is based in Chicago.

Phone number: (312) 281-2470

Email: magda.lasota@gfk.com

GfK Methodology

Introduction

The GfK Group (formerly Knowledge Networks) is passionate about research in marketing, media, health, and social policy. We collaborate closely with client teams throughout the research process, while applying rigor in everything we do. We specialize in innovative online research that consistently gives leaders in business, government, and academia the confidence to make important decisions. GfK delivers affordable, statistically valid online research through KnowledgePanel® and leverages a variety of other assets, such as world-class advanced analytics, an industry-leading physician panel, an innovative platform for measuring online ad effectiveness, and a research-ready behavioral database of frequent supermarket and drug store shoppers.

GfK has recruited the first online research panel that is representative of the entire United States population. Panel members are randomly recruited through probability-based sampling, and households are provided with access to the Internet and hardware if needed.

GfK recruits panel members by using address-based sampling methods [previously GfK relied on random-digit dialing methods]. Once household members are recruited for the panel and assigned to a study sample, they are notified by email for survey taking, or panelists can visit their online member page for survey taking (instead of being contacted by telephone or postal mail). This allows surveys to be fielded very quickly and economically. In addition, this approach reduces the burden placed on respondents, since email notification is less intrusive than telephone calls, and most respondents find answering Web questionnaires more interesting and engaging than being questioned by a telephone interviewer. Furthermore, respondents have the convenience to choose what time of day to complete their assigned survey.

Documentation regarding KnowledgePanel sampling, data collection procedures, weighting, and IRB-bearing issues are available at the below online resources.

- <http://www.knowledgenetworks.com/ganp/reviewer-info.html>
- <http://www.knowledgenetworks.com/knpanel/index.html>
- <http://www.knowledgenetworks.com/ganp/irbsupport/>

The GfK Group

The GfK Group has a strong tradition in working with sophisticated academic, government, and commercial researchers to provide high quality research, samples, and analyses. The larger GfK Group offers the fundamental knowledge for governmental agencies, academics, industries, industry, retailers, services companies and the media need to provide exceptional quality in research to make effective decisions. It delivers a comprehensive range of information and consultancy services. GfK is one of the leading survey research organizations worldwide, operating in more than 100 countries with over 13,000 research staff. In 2013, the GfK Group's sales amounted to EUR 1.49 billion.

For further information, visit our website: www.gfk.com.

KnowledgePanel Methodology Information

Complete and current information about KnowledgePanel sampling and recruitment methodology and design is available at

<http://www.gfk.com/Documents/GfK-KnowledgePanel-Design-Summary.pdf>

KnowledgePanel's recruitment process was originally based exclusively on a national Random Digit Dialing (RDD) sampling methodology. In order to improve the representation of the panel, GfK migrated to using an Address Based Sampling (ABS) methodology for selecting panel members in 2009. This probability-based sampling methodology improves population coverage, and provides a more effective sampling infrastructure for recruitment of hard-to-reach individuals, such as young adults and those from various minority groups. It should be noted that under the ABS recruitment households without Internet connection are provided with a web-enabled device and free Internet service.

After initially accepting the invitation to join the panel, participants are asked to complete a short demographic survey (the initial profile survey); answers to which allow efficient panel sampling and weighting for future surveys. Completion of the profile survey allows participants to become panel members, and as in the past, all respondents are provided the same privacy terms and confidentiality protections.

ABS Recruitment.

Our recruitment protocol relies on probability-based sampling of addresses from the United States Postal Service's Delivery Sequence File (DSF). The key advantage of the ABS methodology is that it allows sampling of almost all United States households. Regardless of household telephone status, all households can be reached and contacted through postal mail. In late 2009 the ABS recruitment included a geographic stratification, whereby Census Block Groups (CBG) with high density minority communities were oversampled (Stratum 1) and the balance of the CBGs (Stratum 2) were slightly under-sampled. In 2010, the two strata were further redefined to target high density Hispanic areas in Stratum 1 and all else in Stratum 2. In 2011, pre-identified ancillary information about addresses was used to construct and target households in the following four sampling strata:

1. Hispanic ages 18-24
2. Non-Hispanic ages 18-24
3. Hispanic ages 25+
4. Non-Hispanic ages 25+

In 2012, a similar four-stratum design was implemented, with the ages changing to 18-29 and 30+ for both the Hispanic and non-Hispanic strata. As detailed below, specific adjustments are applied to compensate for any oversampling that is carried out to improve the demographic composition of the panel.

Randomly sampled addresses from the DSF are invited to join KnowledgePanel through a series of mailings, including an initial invitation letter, a reminder postcard, and a subsequent follow-up letter. Given that approximately 45% of the physical addresses can be matched to a corresponding landline telephone number, about 5 weeks after the initial mailing, telephone refusal-conversion calls are made to households for whom a telephone number was matched to the sampled address. Invited households can join the panel by:

- Completing and mailing back a paper form in a postage-paid envelope
- Calling a toll-free hotline phone number maintained by GfK
- Going to a designated GfK website and completing the recruitment form at the website

Household Member Recruitment.

For all recruitment efforts, during the initial recruitment survey, all household members are enumerated. Following enumeration, attempts are made to recruit every household member who is at least 13 years old to participate in KnowledgePanel surveys. For household members aged 13 to 17, consent is collected from the parents or the legal guardian during the initial recruitment interview. If no consent is given, no further direct communication with the teenagers is attempted.

Survey Sampling from KnowledgePanel

Once panel members are recruited and profiled, they become eligible for selection for client surveys. In most cases, the specific survey sample represents an equal probability selection method (EPSEM) sample from the panel for general population surveys. Customized stratified random sampling based on profile data can also be conducted as required by the study design.

The general sampling rule is to assign no more than one survey per week to individual members. Allowing for rare exceptions during some weeks, this limits a member's total assignments per month to four or six surveys. In certain cases, a survey sample calls for pre-screening, that is, members are drawn from a subsample of the panel (such as females, Republicans, grocery shoppers, etc.). In such cases, care is taken to ensure that all subsequent survey samples drawn that week are selected in such a way as to result in a sample that remains representative of the panel distributions.

For this survey, a nationally representative sample of U.S. adults (18 and older) and a representative sample of pre-identified smokers (every day or some days) were selected.

Survey Administration

Once assigned to a survey, members receive a notification email letting them know there is a new survey available for them to take. This email notification contains a link that sends them to the survey questionnaire. No login name or password is required. The field period depends on the client's needs and can range anywhere from a few hours to several weeks.

After three days, automatic email reminders are sent to all non-responding panel members in the sample. If email reminders do not generate a sufficient response, an automated telephone reminder call can be initiated. The usual protocol is to wait at least three to four days after the email reminder before calling. To assist panel members with their survey taking, each individual has a personalized "home page" that lists all the surveys that were assigned to that member and have yet to be completed.

GfK also operates an ongoing modest incentive program to encourage participation and create member loyalty. Members can enter special raffles or can be entered into special sweepstakes with both cash rewards and other prizes to be won.

The typical survey commitment for panel members is one survey per week or four per month with duration of 10 to 15 minutes per survey. In the case of longer surveys, an additional incentive is typically provided.

Sample Weighting

As detailed above, significant resources and infrastructure are devoted to the recruitment process for the KnowledgePanel (KP) so that the resulting panel can properly represent the adult population of the US. This representation is not only achieved with respect to a broad set of geo-demographic distributions, but also hard-to-reach adults – such as those without landline telephone or Spanish language dominant individuals – are recruited in proper proportions as well. Consequently, the raw distribution of KP mirrors that of the US adults fairly closely, barring occasional disparities that may emerge for certain subgroups due to differential attrition rates among recruited panel members.

For selection of general population samples from KP, however, a patented methodology has been developed that ensures the resulting samples behave as EPSEM. Briefly, this methodology starts by weighting the entire KP to the benchmarks secured from the latest March supplement of the Current Population Survey (CPS) along several dimensions. This way, the weighted distribution of KP perfectly matches that of the US adults – even with respect to the above mentioned few dimensions where minor misalignments may result from differential attrition rates.

Typically, the geo-demographic dimensions used for weighting the entire KP include:

- Gender (Male/Female)
- Age (18–29, 30–44, 45–59, and 60+)
- Race/Hispanic ethnicity (White/Non-Hispanic, Black/Non-Hispanic, Other/Non-Hispanic, 2+ Races/Non-Hispanic, Hispanic)
- Education (Less than High School, High School, Some College, Bachelor and beyond)
- Census Region (Northeast, Midwest, South, West)
- Household income (under \$10k, \$10K to <\$25k, \$25K to <\$50k, \$50K to <\$75k, \$75K to <\$100k, \$100K+)
- Home ownership status (Own, Rent/Other)
- Metropolitan Area (Yes, No)
- Internet Access (Yes, No)

Using the above weights as the measure of size (MOS) for each panel member, in the next step a PPS (probability proportional to size) procedure is used to select study specific samples. It is the application of this PPS methodology with the above MOS values that produces fully self-weighting samples from KP, for which each sample member can carry a design weight of unity. Moreover, in instances where the study design has required any form of oversampling of specific subgroups, such departures from an EPSEM design are corrected by adjusting the corresponding design weights accordingly with the CPS benchmarks serving as reference points.

Study-Specific Post-Stratification Weights

Once the sample has been selected and fielded, and all the study data are collected and made final, a post-stratification weight is computed to adjust for any survey non-response as well as any non-coverage or under- and over-sampling resulting from the study-specific sample design. Demographic and geographic distributions for the non-institutionalized, civilian population ages 18+ from the most recent CPS are used as benchmarks in this adjustment.

The following benchmark distributions are utilized for this post-stratification adjustment for the general population respondents:

- Gender (Male/Female)
- Age (18–29, 30–44, 45–59, and 60+)
- Race/Hispanic ethnicity (White/Non-Hispanic, Black/Non-Hispanic, Other or 2+ Races/Non-Hispanic, Hispanic)
- Education (Less than High School, High School, Some College, Bachelors and higher)
- Household income (under <\$25k, \$25K to <\$50k, \$50K to <\$75k, \$75K+)
- Census Region (Northeast, Midwest, South, West)
- Metropolitan Area (Yes, No)
- Internet Access (Yes, No)

The weighting benchmarks of smokers and non-smokers were derived from the weighted general population respondents. Reconfirmed smokers were then weighted to look like age 18+ smokers on the same demographics used for the general population respondents.

The combined general population and smoker augment were also weighted to look like a mega-general population sample. Demographics were controlled within smokers and non-smokers for the mega-general population sample on the same demographics benchmarks as the general population respondents.

Comparable distributions are calculated by using all completed cases from the field data. Since study sample sizes are typically too small to accommodate a complete cross-tabulation of all the survey variables with the benchmark variables, a raking procedure is used for the post-stratification weighting adjustment. Using the base weight as the starting weight, this procedure adjusts the sample data back to the selected benchmark proportions. Through an iterative convergence process, the weighted sample data are optimally fitted to the marginal distributions.

After this final post-stratification adjustment, the distribution of the calculated weights are examined to identify and, if necessary, trim outliers at the extreme upper and lower tails of the weight distribution. The post-stratified and trimmed weights are then scaled to the sum of the total sample size of all eligible respondents and to respective population sizes.

For Year 2, we also combined the year-specific weights with respect to their effective sample sizes to create two new pooled weights. We weighted year 1 and year 2 respondents to look like 18+ US population on the standard weighting (using the latest CPS data for benchmarks).

Weights Definition:

weight1: Gen Pop Respondents weights [sample-scaled]

weight1pop: Gen Pop Respondents weights [population-scaled]

weight2: Total Respondents (Gen Pop + Pre-Id Smoker Augment) weights [sample-scaled]

weight2pop: Total Respondents (Gen Pop + Pre-Id Smoker Augment) weights [population-scaled]

weight3: Total Smoker Respondents (Gen Pop + Pre-Id Smoker Augment) weights [sample-scaled]

weight3pop: Total Smoker Respondents (Gen Pop + Pre-Id Smoker Augment) weights [population-scaled]

weight4: Total Years 1 and 2 Respondents (Gen Pop + Pre-Id Smoker Augment) weights [sample based]

weight4pop: Total Years 1 and 2 Respondents (Gen Pop + Pre-Id Smoker Augment) weights
[population based]

Trimming:

weight1: (1.03%, 99.01%)

weight1pop: None – Scaled weights-scaled on “weight1”

weight2:

smokers: (2.02%, 98.05%)

non-smokers: None

weight2pop: None – Scaled weights-scaled on “weight2”

weight3: None – Scaled weights-scaled on “weight2”

weight3pop: None – Scaled weights-scaled on “weight2”

weight4: (0.99%, 99.00%)

weight4pop: None – Scaled weights based on “weight4”

Design Effect:

weight1: 1.6154

weight1pop: 1.6154

weight2: 1.6799

weight2pop: 1.6799

weight3: 2.1290

weight3pop: 2.1290

weight4: 1.4408

weight4pop: 1.4408

Valid	0 0 surveys	2557	42.3	42.3	42.3
	1 1 survey	1874	31.0	31.0	73.2
	2 2-5 surveys	1325	21.9	21.9	95.1
	3 6+ surveys	295	4.9	4.9	100.0
	Total	6051	100.0	100.0	

DeviceType DOV: Device Type

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 iPad	406	6.7	6.7	6.7
	2 iPod	8	.1	.1	6.8
	3 iPhone	474	7.8	7.8	14.7
	4 Mac	394	6.5	6.5	21.2
	5 ChromeBook	37	.6	.6	21.8
	6 BlackBerry	4	.1	.1	21.9
	7 AndroidTablet	147	2.4	2.4	24.3
	8 AndroidPhone	658	10.9	10.9	35.2
	9 TabletOther	5	.1	.1	35.3
	10 WinPC	3796	62.7	62.7	98.0
	11 WinPhone	9	.1	.1	98.1
	14 KindleFireTablet	75	1.2	1.2	99.4
	16 others	38	.6	.6	100.0
	Total	6051	100.0	100.0	

Appendix C: KnowledgePanel® Response Rate Report

KnowledgePanel® is a probability-based panel. By definition, all members of KnowledgePanel® have a known probability of selection. As a result, it is mathematically possible to calculate a proper response rate that takes into account all sources of nonresponse. Below are the components of the response rate calculation and the actual calculations. An extended description of how to compute response metrics for online panels can be found in:

Callegaro, Mario & DiSogra, Charles (2008). Computing response metrics for online panels. *Public Opinion Quarterly* 72(5). pp. 1008-1032.²

Response Rate Summary Metrics:

A. Number of Assigned Panelists	8,135
B. Study-Specific Average Panel Recruitment Rate (RECR)	13.8%
C. Study-Specific Average Household Profile Rate (PROR)	64.6%
D. Study-Specific Average Household Retention Rate (RETR)	32.1%
E. Number of Total Study Completes	6,181
F. Study Completion Rate (COMR)	76.0%
G. Number of Study Break-offs	292
H. Study Breakoff Rate (BOR)	4.5%
I. Number of Qualified Completes	6,091
J. Study Qualification Rate (QUALR)	98.5%
K. Cumulative Response Rate	6.8%

Comparison of Response Rates

It is important to note the differences between a Random Digit Dial (RDD) telephone or mail sample and KnowledgePanel®. RDD telephone and mail samples can be compared because they are one-time surveys. However, an online panel such as KnowledgePanel® is composed of people recruited at different times and, more importantly, committed to answering multiple surveys for a period of time and not just a single survey. Further, with KnowledgePanel®, Panelists must also complete profile surveys in order to become members of the Panel. These differences are reflected in the recruitment and profile rates reported above. These differences make directly comparing response rates between one-time surveys and Panel surveys difficult and perhaps not illuminating.

Opt-in web panels do not permit the calculation of a response rate since the probabilities of selection are unknown. Consequently, opt-in panels are only mathematically capable of computing the survey

² The full text of the paper is available on the Public Opinion Quarterly – Special issue webpage: http://www.oxfordjournals.org/our_journals/poq/special.html

completion rate, which represents the final stage of gaining the cooperation of survey research subjects and excludes the nonresponse resulting from panel recruitment, connection, and panel retention. In addition, studies relying on opt-in intercept, sometimes called “river,” samples where respondents are recruited for a particular survey using various banner or pop-up ads placed on numerous websites rather than from a panel, also have no known selection probabilities and are therefore unable to report response rates. Further, such opt-in online intercept studies are unable to compute completion rates since a sample is not selected and can only report survey breakoff rates.

Practical Advice for Reporting Response Rates

Many journals ask for the final stage completion rate that can be easily reported. Break-off rates are also another indicator of quality.

An example of reporting response metrics is the following:

A random sample of 25,288 panel members was drawn from GfK’s KnowledgePanel®. 17,927 (excluding breakoffs) responded to the invitation and 1,253 qualified for the survey, yielding a final stage completion rate of 70.9% and a qualification rate of 7.0% percent. The recruitment rate for this study, reported by GfK, was 16.4% and the profile rate was 63.5%, for a cumulative response rate of 7.4%.

Formulas Used for Calculations

The formulas, from Callegaro & DiSogra (2008), used to calculate the response summary metrics reported above are presented in Appendix A. Respondent-level cohort recruitment, profile, and retention rates are calculated for each study respondent and averaged across all study respondents to yield the study-specific rates reported on the previous page.

Formulas Used for Response Summary Metric Calculations

Respondent-level Panel Recruitment Rate (RECR):

$$= \frac{\text{Initial Consent}}{\text{Initial Consents} + (\text{Refusals} + \text{Noncontacts} + \text{Other Cases}) + e(\text{Unknown if Household Occupied} + \text{Unknown Other})}$$

Respondent-level Profile Rate (PROR):

$$= \frac{(\text{Profile Completes})}{(\text{Profile Complete} + \text{Partial Profile Complete}) + (\text{Profile Refusals} + \text{Profile Noncontacts} + \text{Other Profile Cases})}$$

Respondent-level Retention Rate (RETR):

$$= \frac{(\text{Profile Completes Present at Time of Study} + \text{Partial Profile Completes Present at Time of Study})}{\text{Profile Complete} + \text{Partial Profile Completes}}$$

Breakoff Rate (BOR):

$$= \frac{\text{Break-offs}}{\text{Study Completes} + \text{Study Partial Completes} + \text{Break-offs}}$$

Qualification Rate (QUALR):

$$= \frac{\text{Qualified Study Complete}}{\text{Qualified Study Complete} + \text{Not-qualified Study Complete}}$$

Study Completion Rate (COMR):

$$= \frac{(\text{Study Completes})}{(\text{Study Completes} + \text{Study Partial Completes}) + (\text{Study Refusals} + \text{Study Noncontacts} + \text{Other Study Cases})}$$

Cumulative Response Rate (CUMRR):

$$= RECR * PROR * COMR$$

